Risk Assessment Models for Non-target and Biodiversity Impacts of GMOs
There are many ways to conduct an ecological risk assessment.
Alternative ERA models

- Ecotoxicology model
- Total biodiversity model
- Functional model (ecological)
Overview of Ecotoxicology Model

- Choose surrogate species from a list
- Conduct acute toxicity experiment using 100-1000x environmental concentration (tier 1, effects characterization)
- If no observable effect, then ERA is finished
- If effect is observed, go to next tier and conduct more experiments
Overview of Total Biodiversity model

- Survey all species in an area
- Test as many species as possible
- Conduct field experiments to evaluate effects on biodiversity (no tiers)
- If no observable effect, then monitor after release of GMO
Overview of Functional model

- Identify and select potential adverse effects associated with GM crop (tier 1)
- Choose indicator species from the receiving environment
- Identify and select risk hypotheses involving species (tier 2)
- Develop plan to falsify or quantify potential risk
1. Formulation of the risk assessment problem

- Ecotoxicology – *Average risk* to non-target species
- Total Biodiversity – *All species* of similar concern
- Functional – Focus on species associated with *highest risks*
Many Species – Different Risks

- Total Biodiversity model
- Functional model
  Biggest Concerns

- Ecotoxicology model
  Average Risk

- Likelihood

- Size of Adverse Effect

GMO ERA Project
Critique of Models

- Ecotoxicology model: We are not concerned about the average risk; more concerned about high risks
- Total Biodiversity model: Not all species are of similar value; some species (and their functions) are more valuable
- Functional model: how do we choose indicator species for risk assessment?
Which species do we select to test?

>600 non-target arthropod species

Thousands of species?
2. Selecting species

- Ecotoxicology model – surrogate species concept
- Total Biodiversity model – As many species as possible
- Functional model – indicator species concept
Surrogate species – Ecotoxicology model

- Usually not in the receiving environment
- Presumed sensitive to stressor
- Supposed to represent a typical species in the environment
- Chosen for ease of rearing, uniformity in test organisms
- Requires high quality rearing facilities

1. Avian oral toxicity test – upland game bird, waterfowl species, broiler chicken
2. Wild mammal oral toxicity test – rodent species
3. Freshwater fish oral toxicity test – guppy
4. Freshwater invertebrate testing - *Daphnia*
5. Estuarine and marine animal testing – grass shrimp and fathead minnow
6. Honey bee testing – larval and adult bee toxicity
7. Predatory insect – most commonly *Chrysoperla carnea* larvae and ladybird beetle
8. Parasitic wasp – *Nasonia vitripennis*
9. Soil invertebrates
Parasitoid Surrogate Species

*Nasonia vitripennis* (an idiobiont parasitoid of housefly pupae)
Surrogate species

- Research has demonstrated that surrogate species do **NOT** represent typical species in the environment (Suter 2007)
- Research has demonstrated that surrogate species are **NOT** the most sensitive species (Suter 2007)
Selecting species – Total Biodiversity model

- Species are from receiving environment
- Difficult to choose, so as many as possible are used
- Represent the range of species in the receiving environment
Selecting species – Functional model

- Species are from receiving environment
- Uses MCDA to identify **indicator species** representing the potentially highest risks to the receiving environment
- Focuses effort on the few species of greatest concern
# 3. Causal (Risk) hypotheses

<table>
<thead>
<tr>
<th>Model</th>
<th>Complexity of risk hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecotoxicology model</td>
<td><strong>Only</strong> considers risks associated with non-target mortality</td>
</tr>
<tr>
<td>Total Biodiversity model</td>
<td>Extremely <strong>complex</strong></td>
</tr>
<tr>
<td>Functional model</td>
<td>Considers <strong>mortality</strong> and <strong>non-mortality risks</strong> and <strong>indirect risks</strong> (disease transmission)</td>
</tr>
<tr>
<td>Model</td>
<td>Complexity of Hypotheses</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Ecotoxicology model model</td>
<td><strong>Simplistic</strong></td>
</tr>
<tr>
<td>Total Biodiversity model</td>
<td><strong>Complicated</strong></td>
</tr>
<tr>
<td>Functional model model</td>
<td>Intermediate complexity</td>
</tr>
</tbody>
</table>
## Risk hypotheses

<table>
<thead>
<tr>
<th>Model</th>
<th>Approach to stressors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecostoxicology model</td>
<td><strong>Narrow</strong>, Transgene product (chemical)</td>
</tr>
<tr>
<td>Total Biodiversity model</td>
<td><strong>Broadly considered</strong>, includes GM plant and transgene</td>
</tr>
<tr>
<td>Functional model</td>
<td><strong>Broadly considered</strong>, includes GM plant and transgene</td>
</tr>
</tbody>
</table>
# Risk hypotheses

<table>
<thead>
<tr>
<th>Model</th>
<th>Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecotoxicology model</td>
<td>Requires large quantities of <strong>purified transgene product</strong></td>
</tr>
<tr>
<td>Total Biodiversity model</td>
<td>Many experiments need to be conducted</td>
</tr>
<tr>
<td>Functional model model</td>
<td>Allows a wide range of experimental approaches</td>
</tr>
</tbody>
</table>
## 4. Tiered Assessment

<table>
<thead>
<tr>
<th>Model</th>
<th>Basis of tiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecotox model</td>
<td>Toxicity to surrogates</td>
</tr>
<tr>
<td>Total Biodiversity model</td>
<td>No tiers</td>
</tr>
<tr>
<td>Functional model</td>
<td>Expected likelihood and effect size of risk</td>
</tr>
</tbody>
</table>
Tiered Assessment – Ecotox model

- Requires the use of “safety factors”

Extrapolations from toxicity in surrogate species to all possible effects (endpoints) in all species
Safety Factors – definition

- a multiplicative factor used to allow extrapolation of the results of the study or to account for uncertainty associated with the study. Also known as “margin of safety” or “uncertainty factor”
Safety Factors

Intended to account for

(1) extrapolating from surrogate species to other species *(Entity)*

(2) extrapolating from acute toxicity to all other responses *(Attribute)*

(3) extrapolating from less than the full life cycle to the full life and beyond *(Time)*

Safety Factors have performed poorly!
Tiered Assessment – Total Biodiversity

- Requires many experiments, can be expensive
Tiered Assessment – Functional

- Assumes that the largest effects are assessed, therefore **does not require use of “safety factors”**
- Does not require use of purified transgene products
## Ecotoxicology model

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Has created pressure against broad-spectrum pesticides</td>
<td>• Does not assess actual risks in the receiving environment</td>
</tr>
<tr>
<td>• Simple</td>
<td>• Simplistic</td>
</tr>
<tr>
<td>• Results are repeatable</td>
<td>• High infrastructure cost (labs and facilities)</td>
</tr>
<tr>
<td>• Widely used for pesticide risk assessment</td>
<td></td>
</tr>
<tr>
<td>Advantages</td>
<td>Disadvantages</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>• Very thorough</td>
<td>• Expensive</td>
</tr>
<tr>
<td>• Direct measure of adverse effects</td>
<td>• “Biodiversity” difficult to measure</td>
</tr>
<tr>
<td>• Can be done qualitatively with expert opinion</td>
<td>• Must be done on large areas in the field</td>
</tr>
<tr>
<td>• Can be used at higher tier testing levels</td>
<td>• Qualitative methods are opaque and not easily replicated</td>
</tr>
<tr>
<td>(coupled to another model)</td>
<td></td>
</tr>
</tbody>
</table>
## Functional model

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Focuses effort on greatest risks</td>
<td>• Assessment endpoints must sometimes be developed</td>
</tr>
<tr>
<td>• Species / processes can be tested in lab and field</td>
<td>• Requires well-designed experiments addressing clear risk hypotheses</td>
</tr>
<tr>
<td>• Some methodologies are already available from EcoTox model</td>
<td></td>
</tr>
</tbody>
</table>
Another Story

“Assessing Sustainability”
Ecotoxicology Model

- Promoted by biotechnology industry for GM-crops
ERA Begins

Conduct Laboratory Tests

Observe An Effect
Continue Testing

Observe No Effect

How good are the tests?
Undetected Effects

ERA Done
Natural and Biological Control

- Value high (worldwide US$417 x 10⁹)
- Ecologically sustainable pest control
Exchanges of Scientific Opinion over Ecotoxicology Model

- Shelton et al., 2009a, Trans Res 18: 317-322.
- Shelton et al., 2009b, Environ Entomol 38: 1533-38.
- Continuation?
Some Technical Aspects of Conducting Reviews of the Literature

Meta-analysis: a statistical method for combining data (evidence) across studies

1. Less subjective than summarizing summaries
2. Uses actual data that was statistically significant
3. Uses actual data, including non-significant data
Meta-analysis – Two different null hypotheses

Key assumption: measuring 1 response or many responses?

If 1 response, then the real response is either positive or negative, but not both

Study 1 evidence: significant positive effect
Study 2 evidence: significant negative effect

Combined evidence: **no significant effect**
Method 1 – One Response (Average Effect)

$H_0$: If there is one real response value, then the data (evidence) are distributed standard normal around the real response value, $\mu$. 
If many responses, then the real responses may be positive and negative

Study 1 evidence: significant positive effect
Study 2 evidence: significant negative effect

Combined evidence: possibly a mixture of significant effects

Andow, Lövei & Arpaia, 2009
Method 2 – Many Responses
(Distribution of Effects)

$H_0$: If there are multiple real response values and all of them are zero, then the data (evidence) are distributed standard normal around 0.

Andow, Lövei & Arpaia, 2009
Meta-analysis – Two different null hypotheses

Method 2 – Many Responses (Distribution of Effects)

1. Calculate a measure of distribution
2. Test for normality
3. If non-normal, conclude that there are non-zero effects. Cannot conclude that there is more than one response

Andow, Lövei & Arpaia, 2009
GM-Bt-crops

- Laboratory data
- Natural enemies
- Direct effect of Bt-toxin

- 55 studies
- 273 responses measured
Tests of Normality

- Cochran’s Q (Shelton et al. 2009a, 2009b call this a “heterogeneity test”)
- Fisher’s sum of the log p’s
- Goodness of fit test (Pearson’s Chi-square, log-linear models)
GOF test insensitive to extreme values

The graph illustrates the behavior of the p-value against the extreme value for different GOF tests: Cochran, Fisher, and Log-linear GOF. The p-value is shown on the y-axis, and the extreme value on the x-axis. The p-value is insensitive to extreme values when it remains constant above a certain threshold, which is indicated by the horizontal line p=0.05 in the graph.
Bt-toxins : non-zero direct effects

Finer analysis

- Separate common species
- Separate toxin types
- Separate direct and indirect effects
- Separate response types (survival, development, growth, reproduction, behavior, enzyme activity)
- Controls for non-independence among response types
Test for non-random responses

70% of 66 comparisons $p < 0.05$
20 n.s., average $n = 16.6$
Reported Results

- All of the original papers found **no direct effects**
- All of the review papers found **no direct effects**
  - “Laboratory studies have revealed ... no indication of direct toxic effects [of Bt-toxins].”
- Conclusion: **There are many undetected effects!**
- Assessment method must be improved to assess sustainability accurately
Shelton’s critique was a “way over-reaction”, says an editor at the Entomological Society of America,..., who asked to remain anonymous. “They seem to have read it with eyes predisposed to dismiss anything reflecting poorly on GMOs”
Add Equivalence Tests

- Used to test if the effects are equivalent
- Standard statistics test if the effects are different. Cannot conclude that they are similar.
Equivalence Tests

- Contrasts with standard hypothesis testing
  \[ H_0: \text{Effect A} = \text{Effect B} \]
  \[ H_a: \text{Effect A} \neq \text{Effect B} \]

- Equivalence tests inverts the hypotheses
  \[ H_0: |\text{Effect A} - \text{Effect B}| > \delta \]
  \[ H_a: |\text{Effect A} - \text{Effect B}| \leq \delta \]
Equivalence Tests

H_a

H_0

H_a

A - B

A = B

H_0

H_a

H_0

A - B

A = B

-\delta

0

+\delta
Equivalence tests

- Equivalence standards ($\delta$) must be determined (often ±25%)
- This requires a clear connection between the species tested and the risk under concern
- Can be developed for a Functional model
Conclusions

- Need to improve the ways that governments conduct ERA of key environmental problems
- Many models to conduct ERA
- Choice of model can bias estimate of risk
- Functional models based on modern ecology are more targeted, and accurate
- Equivalence tests can be a useful alternative for statistical tests